

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/758,033DATE: 02/14/97
TIME: 15:14:57

INPUT SET: S15493.raw

This Raw Listing contains the General
Information Section and up to the first 5 pages.

SEQUENCE LISTING

1
2
3 (1) General Information:
4
5 (i) APPLICANT: Clayman, Gary L.
6
7 (ii) TITLE OF INVENTION: Methods and Compositions for the
8 Diagnosis and Treatment of Cancer
9
10 (iii) NUMBER OF SEQUENCES: 14
11
12 (iv) CORRESPONDENCE ADDRESS:
13 (A) ADDRESSEE: Arnold, White and Durkee
14 (B) STREET: P.O. Box 4433
15 (C) CITY: Houston
16 (D) STATE: TX
17 (E) COUNTRY: USA
18 (F) ZIP: 77210-4433
19
20 (v) COMPUTER READABLE FORM:
21 (A) MEDIUM TYPE: Floppy disk
22 (B) COMPUTER: IBM PC compatible
23 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
24 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
25
26 (vi) CURRENT APPLICATION DATA:
27 (A) APPLICATION NUMBER: UNKNOWN
28 (B) FILING DATE: CONCURRENTLY HERewith
29 (C) CLASSIFICATION: UNKNOWN
30
31 (viii) ATTORNEY/AGENT INFORMATION:
32 (A) NAME: Highlander, Steven L.
33 (B) REGISTRATION NUMBER: 37,642
34 (C) REFERENCE/DOCKET NUMBER: INGN:022
35
36 (ix) TELECOMMUNICATION INFORMATION:
37 (A) TELEPHONE: (512) 418-3000
38 (B) TELEFAX: (512) 474-7577
39
40
41 (2) INFORMATION FOR SEQ ID NO:1:
42
43 (i) SEQUENCE CHARACTERISTICS:
44 (A) LENGTH: 2066 base pairs
45 (B) TYPE: nucleic acid
46 (C) STRANDEDNESS: single

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/758,933DATE: 02/14/97
TIME: 15:15:01

INPUT SET: S15493.raw

47 (D) TOPOLOGY: linear

48

49 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

50

| | | |
|----|---|------|
| 51 | CAAAACCTAC CAGGGCAGCT ACGGTTTCCG TCTGGGCTTC TTGCATTCTG GGACAGCCAA | 60 |
| 52 | | |
| 53 | GTCTGTGACT TGCACGTACT CCCCTGCCCT CAACAAGATG TTTTGCCAAC TGGCCAAGAC | 120 |
| 54 | | |
| 55 | CTGCCCTGTG CAGCTGTGGG TTGATTCCAC ACCCCCGCCC GGCACCCGCG TCCGCGCCAT | 180 |
| 56 | | |
| 57 | GGCCATCTAC AAGCAGTCAC AGCACATGAC GGAGGTTGTG AGGCGCTGCC CCCACCATGA | 240 |
| 58 | | |
| 59 | GCGCTGCTCA GATAGCGATG GTCTGGCCCC TCCTCAGCAT CTTATCCGAG TGGAAAGAAA | 300 |
| 60 | | |
| 61 | TTTGCGTGTG GAGTATTTGG ATGACAGAAA CACTTTTCGA CATAGTGTGG TGGTGCCCTA | 360 |
| 62 | | |
| 63 | TGAGCCGCCT GAGGTTGGCT CTGACTGTAC CACCATCCAC TACAACTACA TGTGTAACAG | 420 |
| 64 | | |
| 65 | TTCTTGCATG GGCGGCATGA ACCGGAGGCC CATCCTCACC ATCATCACAC TGGAAAGACTC | 480 |
| 66 | | |
| 67 | CAGTGGTAAT CTA CTGTTGGGAC GGAACAGCTT TGAGGTGCGT GTTTGTGCCT GTCCTGGGAG | 540 |
| 68 | | |
| 69 | AGACCGGCGC ACAGAGGAAG AGAATCTCCG CAAGAAAGGG GAGCCTCACC ACGAGCTGCC | 600 |
| 70 | | |
| 71 | CCCAGGGAGC ACTAAGCGAG CACTGCCCAA CAACACCAGC TCCTCTCCCC AGCCAAAGAA | 660 |
| 72 | | |
| 73 | GAAACCACTG GATGGAGAAT ATTTACCCTC TCAGATCCGT GGGCGTGAGC GCTTCGAGAT | 720 |
| 74 | | |
| 75 | GTTCCGAGAG CTGAATGAGG CCTTGGAAC TCAAGGATGCC CAGGCTGGGA AGGAGCCAGG | 780 |
| 76 | | |
| 77 | GGGGAGCAGG GCTCACTCCA GCCACCTGAA GTCCAAAAAG GGTCACTCTA CCTCCCGCCA | 840 |
| 78 | | |
| 79 | TAAAAAACTC ATGTTCAAGA CAGAAGGGCC TGA CTGAGAC TGACATTCTC CACTTCTTGT | 900 |
| 80 | | |
| 81 | TCCCCACTGA CAGCCTCCCA CCCCATCTC TCCCTCCCCT GCGATTTTGG GTTTTGGGTC | 960 |
| 82 | | |
| 83 | TTTGAACCCT TGCTTGCAAT AGGTGTGCGT CAGAAGCACC CAGGACTTCC ATTTGCTTTG | 1020 |
| 84 | | |
| 85 | TCCCGGGGCT CCACTGAACA AGTTGGCCTG CACTGGTGTT TTGTTGTGGG GAGGAGGATG | 1080 |
| 86 | | |
| 87 | GGGAGTAGGA CATAACAGCT TAGATTTTAA GGTTTTTACT GTGAGGGATG TTTGGGAGAT | 1140 |
| 88 | | |
| 89 | GTAAGAAATG TTCTTGCACT TAAGGGTTAG TTTACAATCA GCCACATTCT AGGTAGGGGC | 1200 |
| 90 | | |
| 91 | CCACTTCACC GTACTAACCA GGGAAGCTGT CCCTCACTGT TGAATTTTCT CTAACCTCAA | 1260 |
| 92 | | |
| 93 | GGCCCATATC TGTGAAATGC TGGCATTTC ACCTACCTCA CAGAGTGCAT TGTGAGGGTT | 1320 |
| 94 | | |
| 95 | AATGAAATAA TGTACATCTG GCCTTGAAAC CACCTTTTAT TACATGGGGT CTAGAACCTG | 1380 |
| 96 | | |
| 97 | ACCCCTTGA GGGTGCTTGT TCCCTCTCCC TGTGGTTCGG TGGGTTGGTA GTTTCTACAG | 1440 |
| 98 | | |
| 99 | TTGGGCAGCT GGTTAGGTAG AGGGAGTTGT CAAGTCTCTG CTGGCCCAGC CAAACCCTGT | 1500 |

INPUT SET: S15493.raw

```

100
101 CTGACAACCT CTTGGTGAAC CTTAGATCCT AAAAGGAAAT GTCACCCCAT CCCACACCCT 1560
102
103 GGAGGATTTT ATCTCTTGTA TAGATGATCT GGATCCACCA AGACTTGTTT TAGCTCAGGG 1620
104
105 TCCAATTTCT TTTTCTTTT TTTTCTTTT TTTCTTTTTC TTTGAGACTG GGTCTCTTTG 1680
106
107 TTGCCCCAGG CTGGAGTGGA GTGGCGTGAT CTGGCTTACT GCAGCCTTTG CCTCCCCGGC 1740
108
109 TCGAGCAGTC CTGCCTCAGC CTCCGGAGTA GCTGGGACCA CAGGTTTCATG CCACCATGGC 1800
110
111 CAGCCAACTT TTGCATGTTT TGTAGAGATG GGGTCTCACA GTGTTGCCCA GGCTGGTCTC 1860
112
113 AAACCTCTGG GCTCAGGCGA TCCACCTGTC TCAGCCTCCC AGAGTGCTGG GATTACAATT 1920
114
115 GTGAGCCACC ACGTCCAGCT GGAAGGGTCA ACATCTTTTA CATTCTGCAA GCACATCTGC 1980
116
117 ATTTTCACCC CACCCTCCC CTCTTCTCCC TTTTATATC CCATTTTAT ATCGATCTCT 2040
118
119 TATTTTACAA TAAAACTTTG CTGCCA 2066
120
121

```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

```

125 (A) LENGTH: 293 amino acids
126 (B) TYPE: amino acid
127 (C) STRANDEDNESS: single
128 (D) TOPOLOGY: linear
129

```

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

130
131
132 Lys Thr Tyr Gln Gly Ser Tyr Gly Phe Arg Leu Gly Phe Leu His Ser
133 1 5 10 15
134
135 Gly Thr Ala Lys Ser Val Thr Cys Thr Tyr Ser Pro Ala Leu Asn Lys
136 20 25 30
137
138 Met Phe Cys Gln Leu Ala Lys Thr Cys Pro Val Gln Leu Trp Val Asp
139 35 40 45
140
141 Ser Thr Pro Pro Pro Gly Thr Arg Val Arg Ala Met Ala Ile Tyr Lys
142 50 55 60
143
144 Gln Ser Gln His Met Thr Glu Val Val Arg Arg Cys Pro His His Glu
145 65 70 75 80
146
147 Arg Cys Ser Asp Ser Asp Gly Leu Ala Pro Pro Gln His Leu Ile Arg
148 85 90 95
149
150 Val Glu Gly Asn Leu Arg Val Glu Tyr Leu Asp Asp Arg Asn Thr Phe
151 100 105 110
152

```

RAW SEQUENCE LISTING

DATE: 02/14/97

PATENT APPLICATION US/08/758,033

TIME: 15:15:07

INPUT SET: S15493.raw

```

153   Arg His Ser Val Val Val Pro Tyr Glu Pro Pro Glu Val Gly Ser Asp
154           115                      120                      125
155
156   Cys Thr Thr Ile His Tyr Asn Tyr Met Cys Asn Ser Ser Cys Met Gly
157           130                      135                      140
158
159   Gly Met Asn Arg Arg Pro Ile Leu Thr Ile Ile Thr Leu Glu Asp Ser
160           145                      150                      155                      160
161
162   Ser Gly Asn Leu Leu Gly Arg Asn Ser Phe Glu Val Arg Val Cys Ala
163           165                      170                      175
164
165   Cys Pro Gly Arg Asp Arg Arg Thr Glu Glu Glu Asn Leu Arg Lys Lys
166           180                      185                      190
167
168   Gly Glu Pro His His Glu Leu Pro Pro Gly Ser Thr Lys Arg Ala Leu
169           195                      200                      205
170
171   Pro Asn Asn Thr Ser Ser Ser Pro Gln Pro Lys Lys Lys Pro Leu Asp
172           210                      215                      220
173
174   Gly Glu Tyr Phe Thr Leu Gln Ile Arg Gly Arg Glu Arg Phe Glu Met
175           225                      230                      235                      240
176
177   Phe Arg Glu Leu Asn Glu Ala Leu Glu Leu Lys Asp Ala Gln Ala Gly
178           245                      250                      255
179
180   Lys Glu Pro Gly Gly Ser Arg Ala His Ser Ser His Leu Lys Ser Lys
181           260                      265                      270
182
183   Lys Gly Gln Ser Thr Ser Arg His Lys Lys Leu Met Phe Lys Thr Glu
184           275                      280                      285
185
186   Gly Pro Asp Ser Asp
187           290
188
189

```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2066 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```

199
200   CAAAACTTAC CAGGGCAACT ATGGCTTCCA CCTGGGCTTC CTGCAGTCTG GGACAGCCAA      60
201
202   GTCTGTTATG TGCACGTACT CTCCTCCCCT CAATAAGCTA TTCTGCCAGC TGGCGAAGAC      120
203
204   GTGCCCTGTG CAGTTGTGGG TCAGCGCCAC ACCTCCAGCT GGGAGCCGTG TCCGCGCCAT      180
205

```

INPUT SET: S15493.raw

| | | | | | | | |
|-----|------------|-------------|-------------|--------------|------------|-------------|------|
| 206 | GGCCATCCAC | AAGAAGTCAC | AGCACTTGAC | GGGGGTCGTG | AGACGCTGCC | CCCACCATGA | 240 |
| 207 | | | | | | | |
| 208 | GCGCTGCTCC | GATGGTGATG | GCCTGGCTCC | TCCCCAGCAT | CTTATCCGGG | TGGAAGGAAA | 300 |
| 209 | | | | | | | |
| 210 | TTTGTATCCC | GAGTATCTGG | AAGACAGGCA | GACTTTTCGC | CACAGCGTGG | TGGTACCTTA | 360 |
| 211 | | | | | | | |
| 212 | TGAGCCACCC | GAGGCCGGCT | CTGAGTATAC | CACCATCCAC | TACAAGTACA | TTTGTAAATAG | 420 |
| 213 | | | | | | | |
| 214 | CTCCTGCATG | GGGGGCATGA | ACCGCCGACC | TATCCTTACC | ATCATCACAC | TGGAAGACTC | 480 |
| 215 | | | | | | | |
| 216 | CAGTGGGAAC | CTTCTGGGAC | GGGACAGCTT | TGAGGTTCGT | GTTTGTGCCT | GCCCTGGGAG | 540 |
| 217 | | | | | | | |
| 218 | AGACCGCCGT | ACAGAAGAAG | AAAATTTCCG | CAAAAAGGAA | GTCCTTTGCC | CTGAAC TGCC | 600 |
| 219 | | | | | | | |
| 220 | CCCAGGGAGC | GCAAAGAGAG | CGCTGCCCAC | CTGCACAAGC | GCCTCTCCCC | CGCAAAAGAA | 660 |
| 221 | | | | | | | |
| 222 | AAAACCACTT | GATGGAGAGT | ATTTACCCCT | CAAGATCCGC | GGGCGTAAAC | GCTTCGAGAT | 720 |
| 223 | | | | | | | |
| 224 | GTTCGGGGAG | CTGAATGAGG | CCTTAGAGTT | AAAGGATGCC | CATGCTACAG | AGGAGTCTGG | 780 |
| 225 | | | | | | | |
| 226 | AGACAGCAGG | GCTCACTCCA | GCTACCTGAA | GACCAAGAAG | GGCCAGTCTA | CTTCCCGCCA | 840 |
| 227 | | | | | | | |
| 228 | TAAAAAACA | ATGGTCAAGA | AAGTGGGGCC | TGACTCAGAC | TGACATTCTC | CAC TTCTTGT | 900 |
| 229 | | | | | | | |
| 230 | TCCCCACTGA | CAGCCTCCCA | CCCCCATCTC | TCCC TCCCCCT | GCCTTTTGGG | TTTTTGGGTCT | 960 |
| 231 | | | | | | | |
| 232 | TTGAACCCCT | GCTTGCAATA | GGTGTGCGTC | AGAAGCACCC | AGGACTTCCA | TTTGCTTTGT | 1020 |
| 233 | | | | | | | |
| 234 | CCCGGGGCTC | CACTGAACAA | GTTGGCCTGC | ACTGGTGTTT | TGTTGTGGGG | AGGAGGATGG | 1080 |
| 235 | | | | | | | |
| 236 | GGAGTAGGAC | ATACCAGCTT | AGATTTTAAAG | GTTTTTACTG | TGAGGGATGT | TTGGGAGATG | 1140 |
| 237 | | | | | | | |
| 238 | TAAGAAATGT | TCTTGCACTT | AAGGGTTAGT | TTACAATCAG | CCACATTCTA | GGTAGGGGCC | 1200 |
| 239 | | | | | | | |
| 240 | CACTTCACCG | TACTAACCAG | GGAAGCTGTC | CCTCACTGTT | GAATTTTCTC | TAAC TTCAAG | 1260 |
| 241 | | | | | | | |
| 242 | GCCCATATCT | GTGAAATGCT | GGCATT TGCA | CCTACCTCAC | AGAGTGCATT | GTGAGGGTTA | 1320 |
| 243 | | | | | | | |
| 244 | ATGAAATAAT | GTACATCTGG | CCTTGAAACC | ACCTTTTATT | ACATGGGGTC | TAGATGACCC | 1380 |
| 245 | | | | | | | |
| 246 | CCTTGAGGTG | CTTGTTCCCT | CTCCCTGTTG | GTCGGTGGGT | TGGTAGTTTC | TACAGTTGGG | 1440 |
| 247 | | | | | | | |
| 248 | CAGCTGGTTA | GGTTGAGGTA | GTTGTCAGGT | CTCTGCTGGC | CCAGCGAAAT | TCTATCCAGC | 1500 |
| 249 | | | | | | | |
| 250 | CAGTTGTTGG | ACCC TGGCAC | CTCAAATGAA | ATCTCACCCCT | ACCCACACC | CTGTAAGATT | 1560 |
| 251 | | | | | | | |
| 252 | CTATCTCTTG | TATAGATGAT | CTGGATCCAC | CAAGACTTGT | TTTAGCTCAG | GGTCCAATTT | 1620 |
| 253 | | | | | | | |
| 254 | CTTTTTTCTT | TTTTTTTTTTT | TTTTTCTTTT | TCTTTGAGAC | TGGGTCTCTT | TGTTGCCCCA | 1680 |
| 255 | | | | | | | |
| 256 | GGCTGGAGTG | GAGTGGCGTG | ATCTGGCTTA | CTGCAGCCTT | TGCCTCCCCG | GCTCGAGCAG | 1740 |
| 257 | | | | | | | |
| 258 | TCCTGCCTCA | GCCTCCGGAG | TAGCTGGGAC | CACAGGTTCA | TGCCACCATG | GCCAGCCAAC | 1800 |

PAGE: 1

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION US/08/758,033

DATE: 02/14/97

TIME: 15:15:14

INPUT SET: S15493.raw

| Line | Error | Original Text |
|------|---------------------------------|---------------------------------|
| 27 | Wrong application Serial Number | (A) APPLICATION NUMBER: UNKNOWN |
| 29 | Wrong Classification | (C) CLASSIFICATION: UNKNOWN |